

Analisis Data Tersensor Interval Dalam Pemodelan Waktu Mendapatkan Pekerjaan Pertama Alumni Universitas Sriwijaya

Alfensi Faruk

Fakultas MIPA, Universitas Sriwijaya
alfensifaruk@unsri.ac.id

Abstrak—Salah satu informasi penting bagi pihak universitas sebagai landasan dalam membuat kebijakan adalah waktu yang dibutuhkan oleh para alumninya untuk mendapatkan pekerjaan pertama setelah kelulusan. Informasi tersebut biasanya diperoleh melalui *tracer study*, yaitu studi peninjauan mengenai situasi terkini seputar pekerjaan para alumni. Pada saat pelaksanaan *tracer study*, waktu mendapatkan pekerjaan pertama biasanya tidak diketahui secara eksak, tetapi hanya berupa interval waktu dimana peristiwa tersebut terjadi. Dalam analisis *survival*, data dengan karakteristik tersebut dikategorikan sebagai data tersensor interval, sehingga metode-metode konvensional dalam analisis *survival* seperti metode tabel kehidupan dan Kaplan-Meier tidak lagi tepat digunakan untuk mengestimasi fungsi *survival*-nya. Tujuan utama dari penelitian ini adalah mengaplikasikan metode-metode analisis *survival* dari data tersensor interval dalam permasalahan waktu untuk mendapatkan pekerjaan pertama. Data diperoleh melalui hasil *tracer study* yang diselenggarakan oleh Universitas Sriwijaya pada tahun 2012. Instrumen pengumpulan data yang digunakan berupa kuesioner yang dapat diakses oleh responden pada situs resmi Universitas Sriwijaya. Terdapat 637 alumni yang ikut berpartisipasi sebagai responden dalam *tracer study* tersebut. Beberapa karakteristik responden yang dilibatkan dalam analisa adalah jenis kelamin, asal fakultas, tahun kelulusan, sumber informasi pekerjaan, pertimbangan memilih pekerjaan, dan pengalaman mengikuti kursus. Untuk mengestimasi fungsi *survival* digunakan metode *non-parametric maximum likelihood estimate*, yang mana berdasarkan hasil perhitungan yang telah dilakukan, peluang tertinggi alumni untuk mendapatkan pekerjaan adalah setelah bulan ke 24 setelah wisuda dan peluang terendah adalah sebelum hari kelulusan. Kemudian, perbandingan fungsi *survival* antar kategori menggunakan uji *logrank* Finkelstein, uji *logrank* Sun, dan uji Wilcoxon-Mann-Whitney. Dalam kasus data waktu tunggu kerja pertama yang diperoleh dalam penelitian ini, pada umumnya hasil perbandingan berdasarkan ketiga jenis pengujian tersebut menghasilkan nilai statistik hitung dan nilai *p-value* yang tidak jauh berbeda, sehingga kesimpulan yang diperoleh mengenai hipotesis awal (H_0) cenderung seragam.

Kata kunci: *tersensor interval, waktu tunggu kerja*

I. PENDAHULUAN

Tracer study atau studi peninjauan, merupakan salah satu cara yang digunakan oleh banyak institusi pendidikan untuk mengetahui situasi kerja terkini dari para alumninya. Salah satu informasi yang biasanya diperoleh melalui *tracer study* adalah waktu untuk mendapatkan pekerjaan pertama dari alumni. Dalam analisis *survival*, waktu tunggu kerja tersebut disebut sebagai waktu *survival*, yaitu waktu yang dimulai dari awal pengamatan hingga terjadinya peristiwa tertentu.

Pada umumnya, informasi tentang waktu tunggu kerja pertama dalam *tracer study* tidak berupa waktu eksak, melainkan hanya berupa interval waktu yang memuat waktu eksak tersebut. Hal ini berarti waktu *survival* tersebut dikategorikan sebagai data *survival* tersensor interval, yaitu suatu kondisi ketika diketahui hanya interval waktu yang memuat kapan peristiwa tersebut terjadi [1]. Data tersensor interval merupakan salah satu tipe dari beberapa bentuk data tersensor, dimana setiap tipe penyensoran tersebut harus dianalisa dengan metode statistika yang berbeda pula [2]. Contoh literatur yang dikhususkan untuk membahas tentang konsep-konsep dan metode analisis dari data tersensor interval dapat dilihat dalam [1].

Penelitian ini bertujuan (i) mengestimasi fungsi *survival* data waktu tunggu kerja pertama yang tersensor interval (ii) membandingkan fungsi-fungsi *survival* antar kategori data tersensor interval. Data *survival* yang digunakan dalam penelitian ini adalah data hasil *tracer study* yang dilaksanakan oleh Universitas Sriwijaya pada tahun 2012. Fungsi *survival* dari data tersebut diestimasi menggunakan metode *nonparametric maximum likelihood estimator* [3], sedangkan perbandingan antar fungsi *survival* dilakukan menggunakan uji *logrank* Finkelstein [4], uji *logrank* Sun [5], dan uji Wilcoxon-Mann-Whitney (WMW) [6]. Hasil penelitian ini bermanfaat untuk menambah kajian ilmiah mengenai penerapan analisis *survival* pada masalah waktu tunggu kerja, serta dapat menjadi bahan pertimbangan bagi pihak perguruan tinggi dalam membuat kebijakan terkait peningkatan daya saing para alumninya di dunia kerja.

II. METODE PENELITIAN

Data dalam penelitian ini diperoleh dari hasil program *tracer study* yang dilaksanakan oleh Universitas Sriwijaya. Pengumpulan data dilaksanakan mulai bulan Januari 2012 hingga bulan Desember 2012 dengan cara meminta para alumni Universitas Sriwijaya untuk mengisi kuesioner yang tersedia di situs resmi universitas (www.unsri.ac.id). Berdasarkan studi tersebut diperoleh 637 orang responden yang telah mengisi kuesioner dan menjadi sampel dalam penelitian ini. Karakteristik dari responden yang dianalisa dalam penelitian ini adalah jenis kelamin, asal fakultas, tahun kelulusan, sumber informasi pekerjaan, pertimbangan memilih pekerjaan, dan pengalaman mengikuti kursus.

III. HASIL DAN PEMBAHASAN

A. Analisis Survival dan Data Tersensor

Analisis *survival* adalah salah satu cabang dari statistika yang mempelajari teknik analisis data *survival*. Data-data yang termasuk ke dalam analisis *survival* paling tidak memuat informasi tentang waktu kejadian, jumlah subjek yang diamati, banyaknya subjek yang mengalami kegagalan, dan banyaknya subjek yang tersensor [7].

Data tersensor adalah data *survival* yang tidak bisa diamati secara utuh. Terdapat dua macam jenis data tersensor, yaitu: (1) sensor titik, yaitu suatu penyensoran yang terjadi pada suatu titik waktu tertentu, yang terdiri atas data tersensor kanan (tipe I, II, dan III), data tersensor kiri, dan data tersensor campuran (kanan dan kiri), (2) sensor interval, yaitu suatu penyensoran yang dilakukan terhadap subjek dengan tidak diketahui secara tepat dimana titik penyensorannya, namun pada interval waktu mana penyensoran terjadi masih dapat diketahui.

B. Fungsi-Fungsi Penting Dalam Analisis Survival

Misalkan T adalah variabel acak positif dan kontinu terhadap waktu *survival* dan $t \geq 0$ adalah waktu amatan yang merupakan jangka waktu terjadinya peristiwa, maka fungsi kepadatan peluang dari t , adalah

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T < t + \Delta t)}{\Delta t}. \quad (1)$$

Fungsi distribusi peluang kumulatif dari variabel acak T , $F(t)$, adalah peluang terjadinya suatu kejadian hingga waktu t yang dapat dituliskan sebagai berikut

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad (2)$$

Fungsi *survival*, $S(t)$, didefinisikan sebagai peluang suatu individu mengalami peristiwa pada saat lebih dari waktu t . Fungsi *survival* ini dapat dituliskan sebagai

$$S(t) = P(T > t) = \int_t^\infty f(x) dx \quad (3)$$

Selain fungsi-fungsi di atas, masih terdapat fungsi penting lainnya dalam analisis *survival*, yaitu fungsi *hazard* $h(t)$. Fungsi *hazard* menyatakan peluang terjadinya suatu peristiwa pada suatu selang waktu yang sangat kecil, yaitu pada interval $(t, t + \Delta t)$, dengan diketahui bahwa individu tersebut belum mengalami kejadian yang diamati sampai pada waktu t . Fungsi *hazard* dapat juga disebut sebagai peluang sesaat terjadinya suatu peristiwa pada waktu t yang dituliskan sebagai

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t+\Delta t) - F(t)}{\Delta t(1-F(t))} = \frac{f(t)}{1-F(t)} = \frac{f(t)}{S(t)}. \quad (4)$$

C. Data dan Karakteristiknya

Dalam tabel 1, terdapat 8 interval dari waktu tunggu mendapatkan pekerjaan pertama responden. Hari kelulusan didefinisikan sebagai waktu awal observasi ($t = 0$), sedangkan saat pengisian kuesioner didefinisikan sebagai waktu akhir observasi. Waktu eksak untuk mendapatkan pekerjaan pertama dari setiap responden tidak diketahui, sehingga waktu *survival* dalam penelitian ini termasuk ke dalam data *survival* tersensor interval (*interval censored*). Interval pertama merupakan rentang waktu bagi individu yang mendapatkan pekerjaan pertama sebelum lulus kuliah, yang disebut juga sebagai data tersensor kiri. Interval ketiga dan keenam adalah rentang waktu bagi individu yang belum mendapatkan pekerjaan pertama pada saat pengisian kuesioner (waktu akhir observasi) dan telah mencari pekerjaan pertama setidaknya 3 dan 12 bulan setelah hari kelulusan. Kedua interval tersebut disebut juga sebagai data tersensor kanan. Begitu juga dengan interval kedelapan, yang merupakan rentang bagi individu yang mendapatkan pekerjaan lebih dari dua tahun juga termasuk sebagai data tersensor kanan. Selain data tersensor interval, terdapat juga data tersensor kiri dan data tersensor kanan, maka waktu *survival* dalam penelitian ini dapat disebut sebagai data tersensor sebarang.

Karakteristik-karakteristik responden dalam tabel 2 merupakan variabel bertipe kategorik, dimana setiap kategori dari masing-masing variabel tersebut diberi kode dengan bilangan asli, sedangkan untuk responden yang tidak menjawab diberi kode *not available* (NA). Jawaban-jawaban yang tidak diisi tersebut disebut sebagai data hilang. Karena jumlah dari data hilang relatif cukup banyak, yakni berjumlah 103 data, maka data perlu diimputasi dengan nilai dugaan sehingga data tersebut menjadi data yang lengkap.

Metode imputasi yang dipilih adalah imputasi *multiple* [8]. Perangkat lunak yang digunakan adalah R, yaitu menggunakan paket *Multivariate Imputation by Chained Equations* (MICE) versi 2.22 [9]. Metode *default* dalam R adalah metode *predictive mean matching*, yang dapat digunakan baik pada variabel bertipe numerik maupun kategorik. Hasil imputasi terhadap data-data hilang tersebut selengkapnya dapat dilihat pada kolom terakhir dalam tabel 2.

TABEL 1. WAKTU MENDAPATKAN PEKERJAAN PERTAMA RESPONDEN

<i>Interval Ke-</i>	<i>Interval Waktu (Dalam Bulan)</i>	<i>Jumlah Responden</i>
1	$(-\infty, 0]$	106
2	$(0, 3]$	222
3	$(0, \infty)$	31
4	$(3, 6]$	117
5	$(6, 12]$	69
6	$(12, \infty)$	13
7	$(12, 24]$	39
8	$(24, \infty)$	40
TOTAL		637

TABEL 2. KARAKTERISTIK RESPONDEN

<i>No.</i>	<i>Karakteristik</i>	<i>Kategori</i>	<i>Kode</i>	<i>Jumlah Data Awal</i>	<i>Jumlah Data Setelah Diimputasi</i>
1	Jenis Kelamin	Laki-laki	1	356	356
		Perempuan	2	281	281
2	Asal Fakultas	Ekonomi	1	24	24
		KIP	2	56	56
		MIPA	3	214	214

		Pertanian	4	97	97
		Sospol	5	107	107
		Teknik	6	71	71
		Ilmu Komputer	7	51	51
		Hukum	8	10	10
		Kedokteran	9	6	6
		Kesehatan Masyarakat	10	1	1
3	Tahun Lulus	> 2002	1	480	480
		<= 2002	2	157	157
4	Sumber Informasi Mendapatkan Pekerjaan	Media Massa	1		197
		Media Sosial	2	131	138
		Dosen	3	51	55
		Keluarga/Relasi	4	149	155
		Langsung	5	60	63
		Perguruan Tinggi	6	27	29
		Tidak Menjawab	NA	37	0
5	Pertimbangan dalam Memilih Pekerjaan	Tantangan Pekerjaan	1	132	141
		Gaji	2	241	251
		Beasiswa	3	54	58
		Dekat Tempat Tinggal	4	32	32
		Benefit	5	45	46
		Lainnya	6	104	109
		Tidak Menjawab	NA	29	0
6	Mengikuti Kursus Sebelum Mendapatkan Pekerjaan	Ya	1	298	310
		Tidak	2	302	327
		Tidak Menjawab	NA	37	0

D. Metode Non-Parametric Maximum Likelihood Estimate (NPMLE)

Dalam bagian ini, diberikan ulasan ringkas mengenai metode NPMLE serta algoritma *self-consistency*. Misalkan suatu studi memuat n buah subjek yang saling bebas dari suatu populasi dengan waktu *survival* $S(t)$ dan T_i melambangkan waktu *survival* dari subjek ke- i , dengan $i = 1, 2, \dots, n$. Kemudian data tersensor interval dari waktu amatan T_i dapat dituliskan sebagai

$$O = \{(L_i, R_i]; i = 1, \dots, n\}, \quad (5)$$

dimana L_i adalah batas kiri dari interval, R_i adalah batas kanan interval, dan $(L_i, R_i]$ adalah data tersensor yang memuat waktu amatan T_i . Selanjutnya, misalkan $\{s_j\}_{j=0}^m$ merupakan notasi untuk elemen-elemen tunggal terurut dari barisan $\{0, L_i, R_i; i = 1, 2, \dots, n\}$, dan didefinisikan $\alpha_{ij} = I(s_j \in (L_i, R_i])$ serta $p_j = S(s_{j-1}) - S(s_j)$, $i = 1, \dots, n$, $j = 1, \dots, m$. Karena p_j adalah fungsi kepadatan peluang [1], maka diperoleh fungsi *likelihood* yang berbentuk

$$L_S(\mathbf{p}) = \prod_{i=1}^n [S(L_i) - S(R_i)] = \prod_{i=1}^n \sum_{j=1}^m \alpha_{ij} p_j, \quad (6)$$

dengan $\mathbf{p} = (p_1, p_2, \dots, p_m)'$.

Terlihat bahwa fungsi *likelihood* (6) hanya bergantung pada fungsi *survival* $S(t)$ yang terletak di dalam barisan $\{S(s_j)\}_{j=1}^m$ dan tidak bergantung pada bagaimana perubahan dari $S(t)$ di antara elemen-elemen s_j . Hal ini berarti, nilai estimator dari $S(t)$ dengan metode NPMLE hanya dapat ditentukan untuk nilai pada s_j saja. Penentuan estimator tersebut dapat dilakukan dengan memaksimumkan (6) yang bersesuaian dengan subjek \mathbf{p} dengan syarat-syarat $\sum_{j=1}^m p_j = 1$ dan $p_j \geq 0$, $j = 1, \dots, m$ [3].

Terdapat beberapa algoritma yang dapat digunakan dalam penentuan estimator-estimator dari $S(t)$ yang memaksimumkan (6), yaitu algoritma *self-consistency* [3], algoritma ICM [11], dan algoritma *hybrid* [12]. Algoritma *self-consistency* merupakan algoritma yang paling sering digunakan oleh para praktisi dan telah tersedia dalam *software* R dengan paket *Icens*. Oleh karena itu, dalam penelitian ini algoritma *self-consistency* dipilih untuk digunakan dalam mengestimasi fungsi-fungsi penting *survival* berdasarkan metode NPMLE. Dalam algoritma *self-consistency*, estimasi dilakukan secara iteratif. Untuk

menurunkan persamaan *self-consistency* dari data tersensor interval, digunakan suatu pendekatan secara langsung berdasarkan estimasi empiris. Secara umum, algoritma ini diawali dengan memperlakukan data tersensor interval sebagai data yang tidak lengkap, yang kemudian dilanjutkan dengan mengaplikasikan algoritma EM (*Expectation-Maximization*). Uraian teknis dari algoritma ini secara lebih lengkap dapat dilihat dalam [1].

Fungsi *survival* $S(t)$ adalah peluang suatu individu tidak mengalami suatu peristiwa tertentu lebih dari waktu t [7]. Hasil estimasi dari $S(t)$ dan fungsi-fungsi penting lainnya menggunakan metode NPMLE dan algoritma *self-consistency* yang dibantu dengan *software* R pada data waktu tunggu kerja pertama alumni Unsri dapat dilihat dalam tabel 3. Sebagai contoh, nilai peluang seorang alumni Unsri belum mendapatkan pekerjaan pertama lebih dari 3 bulan adalah $\hat{S}(2) = 0,8336$. Sedangkan, $\hat{F}(t)$ menyatakan peluang seorang alumni telah mendapatkan pekerjaan pertama sebelum waktu t . Contohnya, $\hat{F}(4)$ menyatakan peluang seorang alumni Unsri telah mendapatkan pekerjaan pertamanya sebelum bulan keenam setelah wisuda adalah sebesar 0,7316.

E. Uji Perbandingan Fungsi Survival

Setiap karakteristik alumni yang ditampilkan dalam tabel 2 merupakan variabel-variabel yang bertipe kategorik. Untuk mengetahui apakah ada perbedaan yang nyata antar fungsi-fungsi *survival* dapat dilakukan dengan dua cara, yaitu menggunakan metode grafik atau menggunakan uji statistik. Perbandingan dengan metode grafik bersifat sangat subjektif karena pengambilan keputusan hanya berdasarkan visual saja, sedangkan hasil berdasarkan uji statistik lebih baik karena menggunakan hukum dan prinsip dalam statistika.

Terdapat tiga uji statistik yang digunakan untuk membandingkan fungsi *survival* dalam penelitian ini, yaitu uji *logrank* Finkelstein, uji *logrank* Sun, dan uji Wilcoxon-Mann-Whitney (WMW). Ketiga statistik uji tersebut diturunkan dari bentuk umum skor uji *logrank* data tersensor interval untuk parameter β yang berbentuk

$$U = \sum_{i=1}^n z_i \left(\frac{\hat{S}'(L_i) - \hat{S}'(R_i)}{\hat{S}(L_i) - \hat{S}(R_i)} \right) \equiv \sum_{i=1}^n z_i c_i \quad (7)$$

dimana $\hat{S}'(L_i)$ adalah turunan pertama terhadap parameter β , dan z_i adalah vektor kovariat berukuran $k \times 1$. Jika z_i adalah vektor dari k kategori, maka baris ke- l dari (7) dapat dituliskan menjadi

$$U_l = \sum_{j=1}^m w_j \left(d_{jl}^* - \frac{n_{jl}^* - d_j^*}{n_j^*} \right), \quad (8)$$

dimana

$$w_j = \frac{\hat{S}(t_j) \hat{S}'(t_{j-1}) - \hat{S}'(t_{j-1}) \hat{S}(t_j)}{\hat{S}(t_j) [\hat{S}'(t_{j-1}) - \hat{S}'(t_j)]}, \quad (9)$$

dan d_{jl}^* adalah ekspektasi jumlah kegagalan di bawah nol dalam interval ke- j untuk kelompok pada kategori l , d_j^* adalah ekspektasi jumlah kegagalan dalam interval ke- j , dan n_{jl}^* serta n_j^* adalah jumlah individu dalam kelompok yang bersesuaian dengan kategorinya. Bentuk eksplisit dari c_i dan w_j untuk masing-masing skor Finkelstein, skor Sun, dan skor WMW, serta penjelasan yang lebih lengkap tentang hubungannya dengan distribusi normal standar (Z) dan distribusi *chi square* (χ^2) dapat dilihat dalam [12], [13], dan [14].

Dalam penelitian ini, perhitungan semua skor dari ketiga uji tersebut dibantu oleh *software* R sehingga diperoleh *output* seperti yang ditulis secara ringkas dalam tabel 3. *Software* R dapat langsung menghitung nilai Z (untuk uji 2 sampel), nilai χ^2 (untuk uji lebih dari 2 sampel), dan p -value dari ketiga uji tersebut menggunakan paket *interval* dan *perm*. Adapun bentuk dari hipotesis uji untuk karakteristik yang memiliki 2 kategori adalah

Ho: Kedua fungsi *survival* sama atau $S_1(t) = S_2(t)$

Ha: Kedua fungsi *survival* tidak sama atau $S_1(t) \neq S_2(t)$,

statistik ujinya adalah statistik Z dengan taraf signifikansi α . Hipotesis awal (Ho) diterima apabila $-Z_{1/2(1-\alpha)} < Z_{stat} < Z_{1/2(1-\alpha)}$ atau jika nilai p -value $> \alpha$. Untuk karakteristik dengan kategori lebih dari dua, digunakan uji k -sampel, yang hipotesis ujinya adalah

Ho: Tidak ada perbedaan antar fungsi *survival* atau $S_1(t) = S_2(t) = \dots = S_k(t)$

Ha: Setidaknya ada satu fungsi *survival* yang berbeda, statistik ujinya adalah *chi square* (χ^2) dengan taraf signifikansi α . Ho diterima jika $\chi^2_{1-\alpha/2} < \chi^2 < \chi^2_{1-\alpha/2}$ atau jika nilai *p-value* $> \alpha$. Hasil perhitungan dari uji perbandingan fungsi-fungsi *survival* untuk 2-sampel ditampilkan dalam tabel 4, sedangkan untuk k-sampel dapat dilihat dalam tabel 5.

TABEL 3. HASIL ESTIMASI FUNGSI-FUNGSI PENTING *SURVIVAL*

Waktu (t)	Selang Waktu (Bulan)	Fungsi Kepadatan Peluang $\hat{f}(t)$	Fungsi Survival $\hat{S}(t)$	Fungsi Distribusi Kumulatif $\hat{F}(t)$	Fungsi Hazard $\hat{h}(t)$
1	$(-\infty, 0]$	0,1664	1	0	0,1664
2	(0,3]	0,3701	0,8336	0,1664	0,4440
3	(3,6]	0,1951	0,4635	0,5365	0,4210
4	(6,12]	0,1150	0,2684	0,7316	0,4285
5	(12,24]	0,0757	0,1534	0,8466	0,4935
6	(24, ∞)	0,0777	0,0777	0,9223	1

TABEL 4. HASIL UJI PERBANDINGAN FUNGSI *SURVIVAL* 2-SAMPEL

Karakteristik	Uji Finkelstein		Uji Sun		Uji WMW		Kesimpulan Ho ($\alpha = 0.05$)
	Z-Stat	P-Value	Z-Stat	P-Value	Z-Stat	P-Value	
Jenis Kelamin	1.88	0.06	1.92	0.06	1.76	0.08	Diterima
Tahun Lulus	3.20	0.00	3.15	0.00	2.48	0.01	Ditolak
Mengikuti Kursus	0.67	0.50	0.69	0.49	0.58	0.57	Diterima

TABEL 5. HASIL UJI PERBANDINGAN FUNGSI *SURVIVAL* K-SAMPEL

Karakteristik	Uji Finkelstein		Uji Sun		Uji WMW		Kesimpulan Ho ($\alpha = 0.05$)
	χ^2	P-Value	χ^2	P-Value	χ^2	P-Value	
Asal Fakultas	24.5	0.00	25.8	0.00	25.6	0.00	Ditolak
Sumber Informasi Pekerjaan	23.6	0.00	24.5	0.00	30.2	0.00	Ditolak
Pertimbangan dalam Memilih Pekerjaan	12.6	0.03	12.9	0.03	9.79	0.08	Ditolak Kecuali Uji WMW

Nilai $z_{1/2(1-\alpha)}$ dengan taraf signifikansi $\alpha = 0.05$ adalah 1.96, sehingga nilai Z-stat dan *p-value* dalam tabel 4 dari karakteristik jenis kelamin dan mengikuti kursus untuk ketiga jenis uji statistik berada pada daerah penerimaan dari Ho. Hal ini berarti, tidak ada perbedaan antara fungsi *survival* individu berjenis kelamin laki-laki dengan perempuan, serta antara fungsi *survival* dari individu yang pernah mengikuti kursus dengan yang tidak mengikuti kursus sebelum mendapatkan pekerjaan pertama. Sedangkan, untuk karakteristik tahun lulus Ho berada pada daerah penolakan, yang berarti fungsi *survival* dari individu yang lulus di bawah tahun 2002 dengan yang lulus pada tahun 2002 atau setelahnya memiliki perbedaan yang nyata. Berdasarkan nilai *p-value* dalam tabel 5, terlihat bahwa hampir semua Ho ditolak kecuali untuk karakteristik pertimbangan memilih pekerjaan dengan uji WMW dimana Ho diterima, yang diakibatkan karena nilai dari *p-value* = 0.08 > 0.05. Hal ini berarti, dengan uji WMW tidak terdapat perbedaan yang signifikan antara fungsi *survival* dari kategori-kategori dalam karakteristik tersebut.

IV. SIMPULAN DAN SARAN

A. Simpulan

Fungsi *hazard* menyatakan peluang seorang alumni mendapatkan pekerjaan pertamanya pada waktu t . Berdasarkan estimasi yang dilakukan, seorang alumni Unsri pasti bekerja jika masa kelulusan lebih dari dua tahun atau pada saat $t=6$. Peluang terendah alumni Unsri untuk mendapatkan pekerjaan pertamanya adalah pada waktu $t = 1$ atau ketika mereka masih kuliah, yaitu sebesar 0,1664. Kemudian, peluang tertinggi alumni Unsri untuk mendapatkan pekerjaan pertama selain pada saat $t = 6$ adalah ketika $t = 5$, yaitu sejak awal bulan ke-12 hingga akhir bulan ke-24 setelah hari kelulusan, tepatnya sebesar 0,4935.

Terdapat tiga jenis uji yang digunakan untuk membandingkan fungsi-fungsi *survival* dalam penelitian ini. Menggunakan taraf signifikansi 5%, kesimpulan yang diperoleh berdasarkan uji Finkelstein dan uji Sun selalu sama. Namun, terdapat satu kasus dimana kesimpulan dari uji Wilcoxon-Mann-Whitney (WMW) berbeda dengan kedua uji lainnya yaitu ketika menguji perbedaan fungsi *survival* dari karakteristik pertimbangan memilih pekerjaan, yang mana diperoleh kesimpulan bahwa tidak ada perbedaan yang signifikan di antara fungsi-fungsi *survival*nya.

B. Saran

Dalam penelitian ini, analisis yang digunakan masih terbatas pada analisis univariat saja. Oleh karena itu, disarankan untuk melanjutkan penelitian ini dengan melakukan analisis multivariat yang bertujuan untuk mengetahui bagaimana pengaruh dari beberapa karakteristik terhadap waktu tunggu kerja pertama alumni.

UCAPAN TERIMA KASIH

Terimakasih sebesar-besarnya diucapkan kepada Lembaga Penelitian (Lemlit) Universitas Sriwijaya, yang telah membiayai penelitian ini melalui skim penelitian Sains Teknologi dan Seni untuk tahun anggaran 2015, sehingga penelitian ini dapat berjalan dengan baik dan diselesaikan tepat pada waktunya.

DAFTAR PUSTAKA

- [1] J. Sun, The statistical analysis of interval-censored failure time data, Newyork: Springer, 2006.
- [2] I. Gijbels, "Censored Data", WIREs Computational Statistics, vol. 2, pp.178-188, April 2010
- [3] B. W. Turnbull, "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data", Journal of the Royal Statistical Society. Series B (Methodological), vol. 38, no. 3, pp. 290-295, 1976
- [4] D. Finkelstein, and R. Wolfe, "A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data", Biometrics, vol. 41, pp. 845-854, 1985
- [5] J. Sun, "A Non-Parametric Test for Interval-Censored Failure Time Data with Application to AIDS Studies", Statistics in Medicine, vol. 15, pp:1387-1395, 1996
- [6] M. P. Fay, "Rank Invariant Tests for Interval Censored Data under the Grouped Continuous Model", Biometrics, vol. 52, pp. 811-822, 1996
- [7] E. T. Lee, and J. W. Wang, Statistical Methods for Survival Data Analysis, 3rd Ed, New Jersey: John Wiley & Sons, 2002
- [8] D.B. Rubin, Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987
- [9] S.V. Buuren, and K. Groothuis-Oudshoorn, "mice: Multivariate Imputation by Chained Equations in R.", Journal of Statistical Software, vol. 45, no. 3, pp. 1-67, 2011
- [10] P. Groeneboom, and J.A. Wellner, Information bounds and non maximum likelihood estimation, New York: Birkhauser, 1992
- [11] J.A. Wellner, and Y. Zhan, "A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data", Journal of the American Statistical Association, vol. 92, pp. 945-959, 1997
- [12] M.P. Fay, "Rank Invariant Tests for Interval Censored Data under the Grouped Continuous Model", Biometrics, vol. 52, pp. 811-822, 1996

- [13] M. P., Fay, “Comparing Several Score Tests for Interval Censored Data”, *Statistics in Medicine*, vol. 18, 273-285, 1999
- [14] M. P. Fay, and P. A. Shaw, “Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data: The interval R package”, *Journal of Statistical Software*, vol. 36, pp. 1-34, 2010